Average Characteristic Polynomials of Determinantal Point Processes

Adrien Hardy [*] [†]

November 29, 2012

Abstract

We investigate the average characteristic polynomial $\mathbb{E}\big[\prod_{i=1}^{N}(z-x_i)\big]$ where the x_i's are real random variables which form a determinantal point process associated to a bounded projection operator. For a subclass of point processes, which contains Orthogonal Polynomial Ensembles and Multiple Orthogonal Polynomial Ensembles, we provide a sufficient condition for its limiting zero distribution to match with the limiting distribution of the random variables, almost surely, as N goes to infinity. Moreover, such a condition turns out to be sufficient to strengthen the mean convergence to the almost sure one for the moments of the empirical measure associated to the determinantal point process, a fact of independent interest. As an application, we obtain from a theorem of Kuijlaars and Van Assche a unified way to describe the almost sure convergence for classical Orthogonal Polynomial Ensembles. As another application, we obtain from Voiculescu's theorems the limiting zero distribution for multiple Hermite and multiple Laguerre polynomials, expressed in terms of free convolutions of classical distributions with atomic measures.

1 Introduction and statement of the results

1.1 Introduction

For any $N \geq 1$, let $x_1, \ldots, x_N$ be a collection of real random variables which forms a determinantal point process associated with a rank N bounded projection operator. This means there exists for each N a Borel measure μ_N on $\mathbb{R}$ and a $\mu_N \otimes \mu_N$-square integrable function $K_N : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that the joint probability distribution on $\mathbb{R}^N$ of $x_1, \ldots, x_N$ reads

$$\frac{1}{N!} \det \Big[K_N(x_i, x_j) \Big]_{i,j=1}^{N} \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i), \tag{1.1}$$

together with the fact that the operator acting on $L^2(\mu_N)$ by

$$\pi_N : f \mapsto \int K_N(x,y) f(y) \mu_N(\mathrm{d}y) \tag{1.2}$$

[*]Institut de Mathématiques de Toulouse, Université de Toulouse, 31062 Toulouse, France.

[†]Department of Mathematics, KU Leuven, Celestijnenlaan 200 B, 3001 Leuven, Belgium. Email address: adrien.hardy@wis.kuleuven.be

is a (non-necessarily orthogonal) projection on an N-dimensional subspace of $L^2(\mu_N)$. Strictly speaking, this is not a standard way to introduce determinantal point processes, but it is easy to obtain from the standard references on the subject [1, 29, 31, 45] that a determinantal point process (in the usual sense) with a kernel satisfying the latter conditions induces such random variables, and vice versa. To these random variables, we associate their average characteristic polynomial,

$$\chi_N(z) = \mathbb{E}\left[\prod_{i=1}^{N}(z - x_i)\right], \qquad z \in \mathbb{C}, \tag{1.3}$$

where the expectation $\mathbb{E}$ refers to (1.1), and we ask the following question: What is a sufficient condition so that the asymptotic distribution of the zeros of χ_N and the limiting distribution of the random variables x_i's coincide as $N \to \infty$? More precisely, if one denotes by $z_1, \ldots, z_N$ the (non-necessarily real nor distinct) zeros of χ_N and introduces the zero counting probability measure

$$\nu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{z_i}, \tag{1.4}$$

the purpose of this work is to investigate the relation between the weak convergence of ν_N and the almost sure weak convergence of the empirical measure of the determinantal point process, namely

$$\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}. \tag{1.5}$$

Let us first observe that if $x_1, \ldots, x_N$ are i.i.d real random variables, say, with law μ_N and (finite) mean m_N, then this relation is trivial. Indeed, note that $\nu_N = \delta_{m_N}$ since the Vieta's formulas yield $\chi_N(z) = (z - m_N)^N$. Thus ν_N and almost surely $\hat{\mu}^N$ converge towards the same limiting probability distribution as $N \to \infty$ if and only if m_N converges to some real number m and $\hat{\mu}^N$ almost surely converges weakly towards δ_m. If μ_N has no atoms, then the x_i's form a determinantal point process, with kernel $K_N(x, y) = \delta_{xy}$ (the Kronecker's delta) and reference measure μ_N, that is associated to the trivial projection on the 0-dimensional subspace of $L^2(\mu_N)$. We emphasize this trivial case does not satisfy our hypotheses and will not be further discussed.

The situation is much more interesting in the case of non-trivial determinantal point processes. It is for example known that the eigenvalues of an $N \times N$ random matrix drawn from the Gaussian Unitary Ensemble (GUE) is a determinantal point process, and that χ_N is the N-th monic (i.e. with leading coefficient one) Hermite polynomial. After an appropriate rescaling, the zero distribution ν_N converges weakly towards the semi-circle distribution as $N \to \infty$, and so is almost surely the spectral measure $\hat{\mu}^N$, although the numerous proofs of these two facts seem quite independent.

In this work we provide a sufficient condition so that, as $N \to \infty$, the convergence of the moments of ν_N is equivalent to the almost sure convergence of the moments of $\hat{\mu}^N$ for a large class of determinantal point processes, see Theorem 1.3. We will

actually show that condition implies the simultaneous moment convergence of ν_N and of the mean distribution $\mathbb{E}\big[\hat{\mu}^N\big]$, defined by $\mathbb{E}\big[\hat{\mu}^N\big](A) = \mathbb{E}\big[\hat{\mu}^N(A)\big]$ for any Borel set $A \subset \mathbb{R}$, and moreover forces the moments of $\hat{\mu}^N$ to concentrate around their means at a rate $N^{1+\epsilon}$, see Theorem 1.8. At this level of generality, the latter concentration result is new and may be of independent interest. To do so, we develop a moment method for determinantal point processes, involving weighted lattice paths, see Section 2.2.

Besides the theoretical aspect, and as we shall illustrate in Section 3 and 4, such a statement provides two useful practical consequences. On the one hand, the almost sure convergence investigation for such determinantal point processes is thus reducible to the asymptotic analysis for the zeros of polynomials, for which analytic tools have been developed in special cases. On the other hand, one can use the probabilistic background of the random models to obtain a description of the limiting zero distribution of average characteristic polynomials, which in particular cases happen to be special functions of interest in other areas of mathematics.

Before stating our results, let us first introduce and discuss what is already known for two important classes of point processes that will be covered by this work.

1.2 Orthogonal Polynomial Ensembles

Examples of Orthogonal Polynomial (OP) Ensembles are provided by eigenvalue distributions of unitary invariant Hermitian random matrices, including the GUE, Wishart and Jacobi matrix models; they also arise from non-intersecting diffusion processes starting and ending at the origin. In the latter examples, μ_N has a density with respect to the Lebesgue measure. They moreover play a key role in the resolution of several problems from asymptotic combinatorics, such as the problem of the longest increasing subsequence of a random permutation, the shape distribution of large Young diagrams, the random tilings of an Aztec diamond (resp. hexagone) with dominos (resp. rhombuses). This time μ_N is a discrete measure. For further information, see [35, 32, 33] and references therein.

The joint distribution of real random variables $x_1, \dots, x_N$ drawn from an OP Ensemble can be written as

$$\frac{1}{Z_N} \prod_{1 \le i < j \le N} \left|x_j - x_i\right|^2 \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i),$$

where Z_N is a positive normalization constant and μ_N is a measure on $\mathbb{R}$ having all its moments. One can rewrite that distribution in the form (1.1) by introducing the symmetric kernel

$$K_N(x, y) = \sum_{k=0}^{N-1} P_{k,N}(x) P_{k,N}(y), \tag{1.6}$$

where $P_{k,N}$ is the k-th orthonormal polynomial for μ_N. This is the kernel associated with the orthogonal projection onto the subspace of $L^2(\mu_N)$ of polynomials having degree at most $N-1$.

An important observation, provided by a classical integral representation for OPs attributed to Heine, see e.g. [16, Proposition 3.8], is that the average characteristic

polynomial χ_N associated to an OP Ensemble equals the N-th monic OP with respect to μ_N. Since OPs are known to have real zeros, ν_N is thus supported on $\mathbb{R}$.

As we shall recall in Section 2, the mean distribution $\mathbb{E}\big[\hat{\mu}^N\big]$ of a determinantal point process equals $\frac{1}{N}K_N(x,x)\mu_N(\mathrm{d}x)$. Quite remarkably, it turns out that when K_N has the form (1.6), the convergence of the mean distribution has been investigated in the approximation theory literature, where it is referred as the weak convergence of the Christoffel-Darboux kernel. Using the determinantal point processes terminology, Nevai [42] and Van Assche [46] actually proved that the simultaneous weak convergence of $\mathbb{E}\big[\hat{\mu}^N\big]$ and ν_N holds for OP Ensembles as soon as a growth condition on the recurrence coefficients of the $P_{k,N}$'s is satisfied (a definition for the recurrence coefficients of OPs is provided in Section 3). Nevertheless, their proofs involve the Gaussian quadrature associated to OPs, an argument which does not seem to be generalizable to more general determinantal point processes. More recently, and in the case where the supports of the measures μ_N are uniformly bounded, Simon [43] proved the simultaneous moment convergence of $\mathbb{E}\big[\hat{\mu}^N\big]$ and ν_N by means of elegant operator-theoretic arguments, which have been of inspiration for this work.

Concerning the almost sure convergence, after ordering the x_i's and z_i's, Dette and Imhof [19] were able to obtain in the case of the GUE, that is the OP Ensemble associated to $\mu_N(\mathrm{d}x) = e^{-Nx^2/2}\mathrm{d}x$, an upper bound for $\mathbb{P}(\max_{i=1}^N |x_i - z_i| > \varepsilon)$, from which the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N follows. They established that the same picture holds for the Wishart matrices, that is for $\mu_N(\mathrm{d}x) = x^{N\alpha}e^{-Nx}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x$ where $\alpha \geq 0$. Dette and Nagel [20] also worked out the Jacobi case, namely $\mu_N(\mathrm{d}x) = (1-x)^{N\alpha}(1+x)^{N\beta}\mathbf{1}_{[-1,1]}(x)\mathrm{d}x$, where $\alpha, \beta > -1$. Their results moreover cover the associated β-Ensembles. In both works, the proofs strongly use the explicit tridiagonal matrix representation for these random matrix models, which is unfortunately not available for more general unitary invariant Hermitian random matrix models nor determinantal point processes.

Let us now describe a larger class of determinantal point process that will be also covered by this work.

1.3 Multiple Orthogonal Polynomial Ensembles

Firstly introduced by Bleher and Kuijlaars [9] to describe the eigenvalue distribution of an additive perturbation of the GUE, breaking the unitary invariance, Multiple Orthogonal Polynomial (MOP) Ensembles show up in several perturbed matrix models [8, 10, 18], in multi-matrix models [38, 22, 23] as well, and in non-intersecting diffusion processes with arbitrary prescribed starting points and ending at the origin [39]. For general presentations, see [36, 37] and the references therein.

The joint distribution of real random variables $x_1, \ldots, x_N$ distributed according to a MOP Ensemble has the following form

$$\frac{1}{Z_{\mathbf{n},N}} \prod_{1\leq i<j\leq N} (x_j - x_i) \det \begin{bmatrix} \Big\{ x_j^{i-1} w_{1,N}(x_j) \Big\}_{i,j=1}^{n_1,\,N} \\ \vdots \\ \Big\{ x_j^{i-1} w_{r,N}(x_j) \Big\}_{i,j=1}^{n_r,\,N} \end{bmatrix} \prod_{i=1}^N \mu_N(\mathrm{d}x_i), \tag{1.7}$$

where μ_N is a measure on $\mathbb{R}$ having all its moments, $Z_{\mathbf{n},N}$ is a normalization constant and the weights $w_{1,N}, \ldots, w_{r,N} \in L^2(\mu_N)$ are such that (1.7) is indeed a probability distribution. The multi-index $\mathbf{n} = (n_1, \ldots, n_r) \in \mathbb{N}^r$ depends on N and satisfies $\sum_{i=1}^r n_i = N$, where throughout this paper we denote $\mathbb{N} = \{0, 1, 2, \ldots\}$. Note that we recover OP Ensembles by taking $r = 1$.

It turns out, see Section 4.5, that there exists a sequence $(P_{k,N})_{k\in\mathbb{N}}$ of monic polynomials with $\deg P_{k,N} = k$, and a sequence $(Q_{k,N})_{k\in\mathbb{N}}$ of (non-necessarily polynomial) $L^2(\mu_N)$-functions which are biorthogonal, that is

$$\int P_{k,N}(x) Q_{m,N}(x) \mu_N(\mathrm{d}x) = \delta_{km}, \qquad k, m \in \mathbb{N}, \tag{1.8}$$

such that we can rewrite (1.7) in the form (1.1) with the kernel

$$K_N(x, y) = \sum_{k=0}^{N-1} P_{k,N}(x) Q_{k,N}(y). \tag{1.9}$$

Kuijlaars [36, Proposition 2.2] established that the average characteristic polynomial χ_N associated to (1.7) is the $\mathbf{n}$-th (type II) MOP associated with the weights $w_{i,N}$, $1 \le i \le r$, and the measure μ_N, see Definition 4.1.

The simultaneous convergence of the empirical measure $\hat{\mu}^N$ and the zero distribution ν_N of the associated MOPs is expected for several MOP Ensembles. It is for example the case for non-intersecting squared Bessel paths with positive starting point and ending at the origin. Indeed, for this MOP Ensemble $\mathbb{E}\big[\hat{\mu}^N\big]$ converges towards a limiting measure described in terms of the solution of a vector equilibrium problem, see [39, Theorem 2.4 and Appendix], and the limit of ν_N benefits from the same description [40]. The same situation holds in the two-matrix model with quartic/quadratic potentials, by combining the works [22] and [21]. For the non-intersecting squared Bessel paths model, which is equivalent to a non-centered complex Wishart matrix model, the almost sure convergence of $\hat{\mu}^N$ towards the solution of the vector equilibrium problem has moreover been obtained from a large deviation principle in [28]. For the two matrix model, to prove a large deviation upper bound involving a rate function associated to a vector equilibrium problem is still an open problem, see [24] for further discussion. For these two determinantal point processes, the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N will be a consequence of what follows, see Remark 1.9.

It is now time to describe the general setting for which our results hold.

1.4 Statement of the results

Consider a collection of real random variables $x_1, \ldots, x_N$ which forms a determinantal point process associated with a rank N bounded projection operator π_N on $L^2(\mu_N)$, with kernel K_N. Under these only assumptions, K_N is defined $\mu_N \otimes \mu_N$-almost everywhere, which is not sufficient to characterize a determinantal point process, i.e. (1.1) is not well defined. We then follow [29, Remark 5] and generalize their approach in the following way. The spectral decomposition for compact operators [44, Theorem I.4] provides two biorthogonal families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$

of $L^2(\mu_N)$, namely which satisfy

$$\int P_{k,N}(x)Q_{m,N}(x)\mu_N(\mathrm{d}x) = \delta_{km}, \qquad 0 \le k, m \le N-1, \tag{1.10}$$

such that the following equality holds in $L^2(\mu_N \otimes \mu_N)$

$$K_N(x,y) = \sum_{k=0}^{N-1} P_{k,N}(x)Q_{k,N}(y). \tag{1.11}$$

We take the right hand-side of (1.11) as our definition for K_N. Note that although the $P_{k,N}$'s and $Q_{k,N}$'s are still only defined μ_N-almost everywhere, the probability distribution (1.1) now reads

$$\frac{1}{N!}\det\Big[P_{k-1,N}(x_i)\Big]_{i,k=1}^{N}\det\Big[Q_{k-1,N}(x_i)\Big]_{i,k=1}^{N}\prod_{i=1}^{N}\mu_N(\mathrm{d}x_i) \tag{1.12}$$

and is properly defined. Thus, our class of determinantal point processes matches with the Biorthogonal Ensembles introduced by Borodin [11], where we emphasize that the $P_{k,N}$'s and the $Q_{k,N}$'s are $L^2(\mu_N)$-functions.

Now, given a sequence of such determinantal point processes indexed by N (the number of particles),

$$\left\{ \Big(P_{k,N}\Big)_{k=0}^{N-1},\ \Big(Q_{k,N}\Big)_{k=0}^{N-1},\ \mu_N \right\}_{N\ge1}, \tag{1.13}$$

we assume moreover the following structural assumption to hold.

Assumption 1.1.

(a) For each N, the two families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ can be completed in two infinite biorthogonal families $\big(P_{k,N}\big)_{k\in\mathbb{N}}$ and $\big(Q_{k,N}\big)_{k\in\mathbb{N}}$ of $L^2(\mu_N)$, that is which satisfy

$$\int P_{k,N}(x)Q_{m,N}(x)\mu_N(\mathrm{d}x) = \delta_{km}, \qquad k, m \in \mathbb{N}. \tag{1.14}$$

(b) There exists a sequence $(\mathfrak{q}_N)_{N\ge1}$ of integers having sub-power growth, that is for every $n \ge 1$,

$$\mathfrak{q}_N = o(N^{1/n}) \qquad \text{as } N \to \infty, \tag{1.15}$$

such that for all $k \in \mathbb{N}$,

$$xP_{k,N}(x) \in \mathrm{Span}\Big(P_{m,N}(x)\Big)_{m=0}^{k+\mathfrak{q}_N}.$$

Remark 1.2. OP and MOP Ensembles both satisfy Assumption 1.1 (with $\mathfrak{q}_N = 1$ for all $N \ge 1$). A class of determinantal point processes which satisfy this assumption but for which $\mathfrak{q}_N$ may grow is provided by mixed-type MOP Ensembles, originally introduced by Daems and Kuijlaars to describe non-intersecting Brownian bridges with arbitrary starting and ending points [15]. Delvaux showed that the average characteristic polynomial χ_N is in this case a mixture of MOPs [17].

Let $\mathbb{P}$ be the probability measure associated to the product probability space $\bigotimes_N(\mathbb{R}^N, \mathbb{P}_N)$, where $(\mathbb{R}^N, \mathbb{P}_N)$ is the probability space induced by (1.12). The central theorem of this work is the following.

Theorem 1.3. *Assume there exists $\varepsilon > 0$ such that for every $n \geq 1$,*

$$\max_{k,m\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon,\,\left|\frac{m}{N}-1\right|\leq\varepsilon}\left|\int xP_{k,N}(x)Q_{m,N}(x)\mu_N(\mathrm{d}x)\right| = o(N^{1/n}) \tag{1.16}$$

as $N \to \infty$. Then, for all $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty}\left|\int x^\ell \hat{\mu}^N(\mathrm{d}x) - \int x^\ell \nu_N(\mathrm{d}x)\right| = 0, \qquad \mathbb{P}\text{-almost surely.} \tag{1.17}$$

In practice, the sub-power growth condition (1.16) may be interpreted as the condition that a strong enough normalization for the x_i's has been performed.

Remark 1.4. Assumption 1.1 (a) and (b) provide together for each N the unique decomposition

$$xP_{k,N}(x) = \sum_{m=0}^{k+\mathfrak{q}_N}\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}P_{m,N}(x), \qquad k \in \mathbb{N}. \tag{1.18}$$

Thus (1.16) is a growth condition for the coefficients lying in a specific window of the infinite matrix (i.e. operator on $\ell^2(\mathbb{N})$) associated to the operator $f(x) \mapsto xf(x)$ acting on $\mathrm{Span}(P_{k,N})_{k\in\mathbb{N}}$.

As announced in the introduction, let us now provide more precise statements concerning the concrete uses of Theorem 1.3. Having in mind that probability measures on $\mathbb{R}$ with compact support are characterized by their moments, the following consequence of Theorem 1.3 may be of use to obtain almost sure convergence results.

Corollary 1.5. *Under the assumption of Theorem 1.3, if there exists a probability measure μ^* on $\mathbb{R}$ characterized by its moments such that for all $\ell \in \mathbb{N}$,*

$$\lim_{N\to\infty}\int x^\ell \nu_N(\mathrm{d}x) = \int x^\ell \mu^*(\mathrm{d}x),$$

then $\mathbb{P}$-almost surely $\hat{\mu}^N$ converges weakly towards μ^* as $N \to \infty$.

As an example of application, we will obtain from a result of Kuijlaars and Van Assche a unified way to describe the almost sure convergence of classical OP Ensembles in Section 3, see Theorem 3.1.

Similarly, when one is interested in the limiting zero distribution of χ_N, the following corollary will be of help.

Corollary 1.6. *Under the assumption of Theorem 1.3, if*

(a) *for all N large enough χ_N has real zeros,*

(b) *there exists a probability measure μ^* on $\mathbb{R}$ characterized by its moments such that for all $\ell \in \mathbb{N}$,*

$$\lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \int x^\ell \mu^*(\mathrm{d}x), \tag{1.19}$$

then ν_N converges weakly towards μ^* as $N \to \infty$.

As an example of application, we will obtain in Section 4 a description for the limiting zero distribution of multiple Hermite and multiple Laguerre polynomials, see Theorems 4.5 and 4.6. At the best knowledge of the author, this is the first time that a description of these zero limiting distributions is provided in such a level of generality.

Remark 1.7. Although it is not hard to see from our proofs that Theorem 1.3 continues to hold for determinantal point processes on $\mathbb{C}$ (with the introduction of complex conjugations where needed), Corollaries 1.5 and 1.6 are not true in the complex setting. Indeed, consider the eigenvalues of an $N \times N$ unitary matrix distributed according to the Haar measure, which are known to form an OP Ensemble on the unit circle with respect to its uniform measure. We have $\chi_N(z) = z^N$, and thus $\nu_N = \delta_0$ for all N, but the spectral measure $\hat{\mu}^N$ is known to converge towards the uniform distribution on the unit circle as $N \to \infty$.

On the road to establish Theorem 1.3, we prove the following inequality which basically allows to extend the mean convergence of the moments of $\hat{\mu}^N$ to the almost sure one.

Theorem 1.8. *Under the assumptions of Theorem 1.3, for every $0 < \alpha < 1$ and any $\ell \in \mathbb{N}$, there exists $C_{\alpha,\ell}$ independent of N such that for all $\delta > 0$,*

$$\mathbb{P}\left(\left| \int x^\ell \hat{\mu}^N(\mathrm{d}x) - \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] \right| > \delta \right) \leq \frac{C_{\alpha,\ell}}{\delta^2 N^{1+\alpha}}. \tag{1.20}$$

If moreover $\mathfrak{q}_N$ and the left-hand side of (1.16) are bounded (seen as sequences of the parameter N), then (1.20) also holds for $\alpha = 1$.

Remark 1.9. Having in mind real OP and MOP Ensembles, let us stress that our results combine nicely with a Deift-Zhou steepest descent analysis. Indeed, it is known for such ensembles that one can represent K_N in terms of the solution of a Riemann-Hilbert problem, see [16] (resp. [36]) for OP (resp. MOP) Ensembles. This, in principle, allows to use the Deift-Zhou steepest descent method, which yields a precise asymptotic description of K_N, and related quantities. In particular, the $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s, which turn out to be the recurrence coefficients of OPs and MOPs, see Sections 3 and 4, can be expressed in terms of the solution of the Riemann-Hilbert problem (see [16, (3.31)], resp. [26, Section 5]) and a control of their growth would follow from that steepest descent analysis. Such an asymptotic analysis also typically provides the locally uniform convergence and tail estimates for K_N as $N \to \infty$, from which would follow (1.19), and where the limiting measure μ^* has in general compact support. In most cases, the zeros of χ_N are real; this

is always true for OPs and also for important subclasses of MOPs, like Angelesco or AT systems. Thus, if one assumes the latter to be true, the combination of a successful Deift-Zhou steepest descent analysis together with Corollary 1.6 and Theorem 1.8 would provide the almost sure weak convergence of the empirical measure $\hat{\mu}^N$, and moreover the weak convergence of the zero distribution ν_N of the (M)OPs towards μ^*, without extra effort. For example, the determinantal point processes associated to the non-intersecting squared Bessel paths and the two-matrix model with quartic/quadratic potential discussed in Section 1.3 both satisfy the latter; the asymptotics of the recurrence coefficients are actually explicitly described in [40, (1.11)] and [21, Theorem 5.2] respectively. An other example of MOP Ensemble where the recurrence coefficients are explicit is provided by [5], in relation with the six-vertex model.

The rest of this work is structured as follows. In Section 2, we establish Theorems 1.3 and 1.8. In Section 3, by combining Corollary 1.5 and a result concerning the zero convergence of OPs obtained by Kuijlaars and Van Assche, we provide a unified description for the almost sure convergence of classical OP Ensembles. In Section 4, after a quick introduction to MOPs, we use Corollary 1.6 and Voiculescu's theorems in order to identify the limiting zero distribution of the multiple Hermite and multiple Laguerre polynomials in terms of free convolutions, and moreover derive algebraic equations for their Cauchy-Stieltjes transform.

2 Proof of the main theorems

In a first step to establish Theorems 1.3 and 1.8, we express all the quantities of interest in terms of traces of appropriate operators.

2.1 Step 1 : Tracial representations

Consider a determinantal point process associated to the rank N bounded projector π_N acting on $L^2(\mu_N)$ with kernel K_N given by (1.11), so that

$$\mathrm{Im}(\pi_N) = \mathrm{Span}\Big(P_{k,N}\Big)_{k=0}^{N-1}, \qquad \mathrm{Ker}(\pi_N)^{\perp} = \mathrm{Span}\Big(Q_{k,N}\Big)_{k=0}^{N-1}.$$

The usual definition of a determinantal point process, see e.g. [31], provides for any $n \geq 1$ and any Borel function $f : \mathbb{R}^n \to \mathbb{R}$ the identity

$$\mathbb{E}\Bigg[\sum_{i_1 \neq \cdots \neq i_n} f(x_{i_1}, \ldots, x_{i_n})\Bigg] = \int f(x_1, \ldots, x_n) \det\Big[K_N(x_i, x_j)\Big]_{i,j=1}^{n} \prod_{i=1}^{n} \mu_N(\mathrm{d}x_i), \quad (2.1)$$

where the summation concerns all pairwise distinct indices taken from $\{1, \ldots, N\}$.

Let M be the operator acting on $L^2(\mu_N)$ by

$$Mf(x) = xf(x). \quad (2.2)$$

Then the following holds.

Lemma 2.1. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N}\mathrm{Tr}\big(\pi_N M^\ell \pi_N\big).$$

Proof. By using (2.1) with $n = 1$, (1.11) and the biorthogonality relations (1.10), we obtain

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^N x_i^\ell\right] &= \sum_{k=0}^{N-1}\int x^\ell P_{k,N}(x)Q_{k,N}(x)\mu_N(\mathrm{d}x)\\ &= \sum_{k=0}^{N-1}\langle(\pi_N M^\ell \pi_N)P_{k,N}, Q_{k,N}\rangle_{L^2(\mu_N)}\\ &= \mathrm{Tr}\left(\pi_N M^\ell \pi_N\right).\end{aligned}$$

□

We also represent the variance of the moments in a similar fashion.

Lemma 2.2. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{V}ar\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N^2}\Big(\mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) - \mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big)\Big).$$

Proof. We write

$$\mathbb{V}ar\left[\sum_{i=1}^N x_i^\ell\right] = \mathbb{E}\left[\sum_{i=1}^N x_i^{2\ell}\right] + \mathbb{E}\left[\sum_{i\neq j} x_i^\ell x_j^\ell\right] - \left(\mathbb{E}\left[\sum_{i=1}^N x_i^\ell\right]\right)^2$$

in order to obtain, thanks to (2.1) with $n = 2$ and Lemma 2.1,

$$\mathbb{V}ar\left[\sum_{i=1}^N x_i^\ell\right] = \mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) - \iint x^\ell y^\ell K_N(x,y)K_N(y,x)\mu_N(\mathrm{d}x)\mu_N(\mathrm{d}y).$$

Finally, observe that

$$\begin{aligned}&\iint x^\ell y^\ell K_N(x,y)K_N(y,x)\mu_N(\mathrm{d}x)\mu_N(\mathrm{d}y)\\ =\;& \sum_{k=0}^{N-1}\int x^\ell\left(\int K_N(x,y)y^\ell P_{k,N}(y)\mu_N(\mathrm{d}y)\right)Q_{k,N}(x)\mu_N(\mathrm{d}x)\\ =\;& \sum_{k=0}^{N-1}\langle \pi_N M^\ell \pi_N M^\ell \pi_N P_{k,N}, Q_{k,N}\rangle_{L^2(\mu_N)}\\ =\;& \mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big)\end{aligned}$$

to complete the proof. □

We now check that the average characteristic polynomial χ_N equals the characteristic polynomial of the operator $\pi_N M \pi_N$ acting on $\mathrm{Im}(\pi_N)$.

Proposition 2.3. *If* det *stands for the determinant of endomorphisms of* $\mathrm{Im}(\pi_N)$, *then*

$$\chi_N(z) = \det\big(z - \pi_N M \pi_N\big), \qquad z \in \mathbb{C}.$$

Proof. On the one hand, Vieta's formulas provide

$$\mathbb{E}\left[\prod_{i=1}^{N}(z - x_i)\right] = z^N + \sum_{n=1}^{N}\frac{1}{n!}(-1)^n z^{N-n}\,\mathbb{E}\left[\sum_{i_1\neq\cdots\neq i_n} x_{i_1}\cdots x_{i_n}\right]$$

and (2.1) provides for any $1 \le n \le N$

$$\mathbb{E}\left[\sum_{i_1\neq\cdots\neq i_n} x_{i_1}\cdots x_{i_n}\right] = \int \det\Big[x_j K(x_i, x_j)\Big]_{i,j=1}^{n}\prod_{i=1}^{n}\mu_N(\mathrm{d}x_i).$$

On the other hand, since $\pi_N M \pi_N$ is an integral operator acting on $\mathrm{Im}(\pi_N)$ with kernel $(x, y) \mapsto yK_N(x, y)$, the Fredholm's expansion, see e.g. [27], reads

$$\det\big(z - \pi_N M \pi_N\big) = z^N + \sum_{n=1}^{N}\frac{1}{n!}(-1)^n z^{N-n}\int \det\Big[x_j K_N(x_i, x_j)\Big]_{i,j=1}^{n}\prod_{i=1}^{n}\mu_N(\mathrm{d}x_i),$$

from which Proposition 2.3 follows. □

The next immediate corollary will be of important use in what follows.

Corollary 2.4. *For any* $\ell \in \mathbb{N}$,

$$\int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\mathrm{Tr}\big((\pi_N M \pi_N)^\ell\big).$$

The second step is to rewrite the traces in terms of weighted lattice paths.

2.2 Step 2 : Lattice paths representations

We introduce for each N the oriented graph $\mathcal{G}_N = (\mathcal{V}_N, \mathcal{E}_N)$ having $\mathcal{V}_N = \mathbb{N}^2$ for vertices and for edges

$$\mathcal{E}_N = \Big\{(n, k) \to (n+1, m), \qquad n, k \in \mathbb{N}, \quad 0 \le m \le k + \mathfrak{q}_N\Big\}.$$

To each edge is associated a weight

$$w_N\Big((n, k) \to (n+1, m)\Big) = \langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)},$$

and the weight of a finite length oriented path γ on $\mathcal{G}_N$ is defined as the product of the weights of the edges contained in γ, namely

$$w_N(\gamma) = \prod_{e\in\mathcal{E}_N:\, e\subset\gamma} w_N(e). \tag{2.3}$$

Then the following holds.

Lemma 2.5. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k)} w_N(\gamma), \tag{2.4}$$

where the rightmost summation concerns all the oriented paths on $\mathcal{G}_N$ starting from $(0,k)$ and ending at (ℓ,k).

Proof. It follows inductively on ℓ from (1.18) and the definition (2.3) that

$$(\pi_N M^\ell \pi_N)P_{k,N} = \sum_{m=0}^{N-1}\left(\sum_{\gamma:(0,k)\to(\ell,m)} w_N(\gamma)\right) P_{m,N}, \qquad \ell,k \in \mathbb{N}. \tag{2.5}$$

Thus, we obtain from the biorthogonality relations (1.10)

$$\begin{aligned} \mathrm{Tr}\big(\pi_N M^\ell \pi_N\big) &= \sum_{k=0}^{N-1} \langle (\pi_N M^\ell \pi_N)P_{k,N}, Q_{k,N}\rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k)} w_N(\gamma), \end{aligned} \tag{2.6}$$

and Lemma 2.5 follows from Lemma 2.1.

□

Next, we introduce

$$D_N = \Big\{(n,m) \in \mathbb{N}^2 : \ m \geq N\Big\} \tag{2.7}$$

and obtain a similar representation for the moments of ν_N.

Lemma 2.6. *For any $\ell \in \mathbb{N}$,*

$$\int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N=\emptyset} w_N(\gamma). \tag{2.8}$$

Proof. Similarly than for (2.5), we have

$$(\underbrace{\pi_N M \cdots \pi_N M}_{\ell}\pi_N)P_{k,N} = \sum_{m=0}^{N-1}\left(\sum_{\gamma:(0,k)\to(\ell,m),\ \gamma\cap D_N=\emptyset} w_N(\gamma)\right) P_{m,N}, \qquad \ell,k \in \mathbb{N}. \tag{2.9}$$

Since $\pi_N^2 = \pi_N$, this yields

$$\begin{aligned} \mathrm{Tr}\big((\pi_N M \pi_N)^\ell\big) &= \sum_{k=0}^{N-1} \langle (\underbrace{\pi_N M \cdots \pi_N M}_{\ell}\pi_N)P_{k,N}, Q_{k,N}\rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N=\emptyset} w_N(\gamma) \end{aligned} \tag{2.10}$$

and thus Lemma 2.6, because of Corollary 2.4. □

If we denote by $\gamma(m)$ the ordinate of a path γ at abscissa m, then we can represent the variance of the moments of $\hat{\mu}^N$ in a similar fashion.

Lemma 2.7. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{V}ar\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N^2}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(2\ell,k),\ \gamma(\ell)\geq N} w_N(\gamma). \tag{2.11}$$

Proof. We have already shown in (2.6) that

$$\mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) = \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(2\ell,k)} w_N(\gamma). \tag{2.12}$$

Since

$$(\pi_N M^\ell \pi_N M^\ell \pi_N)P_{k,N} = \sum_{m=0}^{N-1}\left(\sum_{\gamma:(0,k)\to(\ell,m),\ \gamma(\ell)<N} w_N(\gamma)\right)P_{m,N}, \qquad \ell, k \in \mathbb{N},$$

we moreover obtain

$$\mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big) = \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(2\ell,k),\ \gamma(\ell)<N} w_N(\gamma). \tag{2.13}$$

Lemma 2.7 is then a consequence of Lemma 2.2 and (2.12)–(2.13). ◻

We are now in position to complete the proofs of Theorems 1.3 and 1.8.

2.3 Step 3 : Majorations and conclusions

Let us first provide a proof for Theorem 1.3 assuming that Theorem 1.8 holds.

Proof of Theorem 1.3. It is enough to prove that for any given $\ell \in \mathbb{N}$

$$\lim_{N\to\infty}\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right| = 0, \tag{2.14}$$

since (1.17) would then follow from Theorem 1.8 and the Borel-Cantelli lemma. As a consequence of Lemmas 2.5 and 2.6, we obtain

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N\neq\emptyset} w_N(\gamma). \tag{2.15}$$

Since by following an edge of $\mathcal{G}_N$ one increases the ordinate by at most $\mathfrak{q}_N$, the rightmost sum of (2.15) will bring null contribution if k is strictly less that $N-\mathfrak{q}_N\ell$. Observe moreover that the vertices explored by any path γ going from $(0,k)$ to (ℓ,k) for some $N-\mathfrak{q}_N\ell \leq k \leq N-1$ such that $\gamma \cap D_N \neq \emptyset$ form a subset of

$$\Big\{(n,m)\in\mathbb{N}^2: \quad 0\leq n\leq \ell, \quad N-\mathfrak{q}_N\ell \leq m < N+\mathfrak{q}_N\ell\Big\}.$$

As a consequence, if one roughly bounds from above the number of such paths by $(2\mathfrak{q}_N\ell)^\ell$, one obtains from (2.15) that

$$\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right| \\ \leq \frac{(2\mathfrak{q}_N\ell)^\ell}{N} \max_{k,m\in\mathbb{N}:\, \left|\frac{k}{N}-1\right|\leq\frac{\mathfrak{q}_N\ell}{N},\, \left|\frac{m}{N}-1\right|\leq\frac{\mathfrak{q}_N\ell}{N}} \left|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\right|^\ell. \quad (2.16)$$

It then follows from (2.16) together with the growth assumptions (1.15) and (1.16) that (2.14) holds, and the proof of Theorem 1.3 is therefore complete up to the proof of Theorem 1.8. □

We now prove Theorem 1.8 by using similar arguments than in the proof of Theorem 1.3.

Proof of Theorem 1.8. Again, because following an edge of $\mathcal{G}_N$ increases the ordinate of at most $\mathfrak{q}_N$, the rightmost sum of (2.11) brings zero contribution except when $k \geq N - \mathfrak{q}_N\ell$. Observe also that the vertices explored by any path γ going from $(0,k)$ to $(2\ell, k)$ for some $N - \mathfrak{q}_N\ell \leq k \leq N-1$ and satisfying $\gamma(\ell) \geq N$ form a subset of

$$\Big\{(n,m)\in\mathbb{N}^2 : \quad 0\leq n\leq 2\ell, \quad N-2\mathfrak{q}_N\ell \leq m < N+2\mathfrak{q}_N\ell\Big\}.$$

As a consequence, we obtain from Lemma 2.7 the (rough) upper-bound

$$\mathbb{V}ar\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] \\ \leq \frac{(4\mathfrak{q}_N\ell)^{2\ell}}{N^2} \max_{k,m\in\mathbb{N}:\, \left|\frac{k}{N}-1\right|\leq\frac{2\mathfrak{q}_N\ell}{N},\, \left|\frac{m}{N}-1\right|\leq\frac{2\mathfrak{q}_N\ell}{N}} \left|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\right|^{2\ell}. \quad (2.17)$$

Using the sub-power growth/boundedness assumptions on $\mathfrak{q}_N$ and on the left-hand side of (1.16), Theorem 1.8 then follows from (2.17) and the Chebyshev inequality. □

3 Applications to OP Ensembles

Consider a sequence of real OP Ensembles, introduced in Section 1.2, namely a sequence of determinantal point processes of the type (1.13) where the measures μ_N on $\mathbb{R}$ have infinite support, have all their moments, and where $P_{k,N} = Q_{k,N}$ stands for the k-th orthonormal polynomial with respect to μ_N. We recall that it satisfies Assumption 1.1 with $\mathfrak{q}_N = 1$ and that χ_N and $P_{N,N}$ have the same zeros. For every N, the celebrated three term recurrence relation for orthonormal polynomials reads

$$\begin{aligned} xP_{k,N}(x) &= a_{k+1,N}P_{k+1,N}(x) + b_{k,N}P_{k,N}(x) + a_{k,N}P_{k-1,N}(x), \qquad k\geq 1, \\ xP_{0,N}(x) &= a_{1,N}P_{1,N}(x) + b_{0,N}P_{0,N}(x), \end{aligned} \quad (3.1)$$

where $a_{k,N} > 0$ and $b_{k,N} \in \mathbb{R}$ are called the recurrence coefficients. By comparing (3.1) with the (unique) decomposition (1.18), one understands that the hypothesis (1.16) of Theorem 1.3 transposes to a sub-power growth condition as $N \to \infty$ for

$$\max_{k\in\mathbb{N}\,:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} |a_{k,N}|\,, \qquad \max_{k\in\mathbb{N}\,:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} |b_{k,N}|\,. \tag{3.2}$$

Here we shall combine our results with the work [41], where Kuijlaars and Van Assche obtain an explicit formula for the limiting zero distribution of general OPs for which the recurrence coefficients converge to some limit. More precisely, let us use the notation $\lim_{k/N\to s} c_{k,N} = c(s)$ when, for every sequences $(k_n)_{n\in\mathbb{N}}$ and $(N_n)_{n\in\mathbb{N}}$ such that $k_n, N_n \to \infty$ and $k_n/N_n \to s$ as $n \to \infty$, we have $\lim_{n\to\infty} c_{k_n,N_n} = c(s)$. Let us also introduce the equilibrium measure of the interval $[\alpha,\beta]$,

$$w_{[\alpha,\beta]}(\mathrm{d}x) = \begin{cases} \dfrac{1}{\pi}\dfrac{\mathbf{1}_{[\alpha,\beta]}(x)\mathrm{d}x}{\sqrt{(\beta-x)(x-\alpha)}} & \text{for } \alpha<\beta, \\ \delta_\alpha & \text{if } \ \alpha=\beta. \end{cases} \tag{3.3}$$

Then the following holds.

Theorem 3.1. *Assume there exists $\varepsilon > 0$ and two continuous functions $a(s)$, $b(s)$ on $[0, 1+\varepsilon)$ such that*

$$\lim_{k/N\to s} a_{k,N} = a(s), \qquad \lim_{k/N\to s} b_{k,N} = b(s), \qquad s \in [0,1+\varepsilon). \tag{3.4}$$

Then, $\mathbb{P}$-almost surely, the empirical measure $\hat{\mu}^N$ associated to the sequence of OP Ensembles converges weakly towards the probability measure

$$\int_0^1 w_{\left[b(s)-2a(s)\,,\,b(s)+2a(s)\right]}(\mathrm{d}x)\mathrm{d}s. \tag{3.5}$$

The latter measure is defined in the obvious way, that is it evaluates any bounded and continuous function f on $\mathbb{R}$ to

$$\int_0^1 \left(\int f(x) w_{\left[b(s)-2a(s)\,,\,b(s)+2a(s)\right]}(\mathrm{d}x) \right) \mathrm{d}s.$$

As we shall observe below on several examples, all the OP Ensembles associated to classical OPs satisfy the conditions of Theorem 3.1, and one can recover the classical limiting distributions (semi-circle law, Marchenko-Pastur law, arcsine law ...) from the formula (3.5). It is in that sense we claimed Theorem 3.1 provides a unified way to describe the almost sure convergence for classical OP Ensembles.

Proof of Theorem 3.1. [41, Theorem 1.4] provides the weak convergence of ν_N towards (3.5). Let us show that ν_N moreover converges to (3.5) in moments. Indeed, one can see from (3.1) that the spectral radius ρ_N of $\pi_N M \pi_N$, which is the one of the matrix

$$\left[\langle xP_{k,N}, P_{m,N}\rangle_{L^2(\mu_N)}\right]_{k,m=0}^{N-1},$$

satisfies

$$\rho_N \le 2 \sup_{k\le N} |a_{k,N}| + \sup_{k\le N} |b_{k,N}|.$$

Then, (3.4) yields $\sup_N \rho_N < +\infty$, and Proposition 2.3 implies that the supports of the ν_N's are uniformly bounded, from which the convergence in moments of ν_N follows. Since (3.4) moreover provides that the sequences (3.2) are bounded, and because (3.5) has compact support, Theorem 3.1 follows from Corollary 1.5. $\square$

We now provide a (non-exhaustive) list of recurrence coefficients for several rescaled classical OPs, from which one can check that Theorem 3.1 applies. The following formulas are easily obtained from the OP literature, see e.g. [34, Section 9], and obvious change of variables.

	Orthogonal polynomial	$\mu_N(\mathrm{d}x)$	Parameters
1	Hermite	$e^{-Nx^2/2}\mathrm{d}x$	none
2	Laguerre	$x^{N\alpha}e^{-Nx}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x$	$\alpha > -1$
3	Jacobi	$(1-x)^{N\alpha}(1+x)^{N\beta}\mathbf{1}_{[-1,1]}(x)\mathrm{d}x$	$\alpha, \beta > 0$
4	Charlier	$\sum_{x\in\mathbb{N}} \frac{(N\alpha)^x}{x!}\delta_{x/N}$	$\alpha > 0$
5	Meixner	$\sum_{x\in\mathbb{N}} \binom{N\beta+x-1}{x}\alpha^x\delta_{x/N}$	$\alpha \in (0,1),\ \beta > 0$

	OP Ensemble	$(a_{k,N})^2$	$b_{k,N}$
1	GUE	$\frac{k}{N}$	0
2	Wishart	$\frac{k}{N}(\frac{k}{N}+\alpha)$	$\frac{2k+1}{N}+\alpha$
3	Random projections	$\frac{4\frac{k}{N}(\frac{k}{N}+\alpha)(\frac{k}{N}+\beta)(\frac{k}{N}+\alpha+\beta)}{(2\frac{k}{N}+\alpha+\beta)^2((2\frac{k}{N}+\alpha+\beta)^2-\frac{1}{N^2})}$	$\frac{\beta^2-\alpha^2}{(2\frac{k}{N}+\alpha+\beta)(2\frac{k+1}{N}+\alpha+\beta)}$
4	Longest increasing subsequence	$\alpha\frac{k}{N}$	$\alpha+\frac{k}{N}$
5	Random Young diagrams	$\frac{1}{(1-\alpha)^2}\frac{k}{N}(\frac{k}{N}+\beta-\frac{1}{N})$	$\frac{1}{1-\alpha}(\frac{k}{N}+\alpha(\frac{k}{N}+\beta))$

Note that these recurrence coefficients may still satisfy (3.4) if one lets the parameters α, β depend on N in an appropriate way; e.g. letting α or β going to zero as $N \to \infty$.

For more details concerning these OP Ensembles, we refer to [13] for the connection between the product of random projections and Jacobi polynomials, to [33] for the problem of the longest increasing subsequence and Charlier polynomials, and to [32, 12] for the random Young diagrams and Meixner polynomials.

It is finally easy to recover from Theorem 3.1 and the latter list of recurrence coefficients the almost sure convergence results for classical OP Ensembles. See also [41] for further computational examples.

4 Application to multiple orthogonal polynomials

MOPs have been introduced in the context of the Hermite-Padé approximation of Stieltjes functions, which was itself first motivated by number theory after Hermite's proof of the transcendence of e, or Apéry's proof of the irrationality of $\zeta(2)$ and $\zeta(3)$, see [47] for a survey. For our purpose here, we will focus on the so-called type II MOPs, for which the zeros are of important interest since they are the poles of the rational approximants provided by the Hermite-Padé theory. These polynomials generalize orthogonal polynomials in the sense that we consider more than one measure of orthogonalization, and a class of classical MOPs such as multiple versions of the Hermite, Laguerre, Jacobi, Charlier, Meixner, etc, polynomials emerged [14, 4, 3]. They are already the subject of many works where they are studied as special functions; we refer to the monograph [30] for further information.

It turns out that even for the multiple Hermite or multiple Laguerre polynomials, no general description of the limiting zero distribution seems yet available in the literature. Our purpose in this section is to obtain such descriptions as a combination of our results with ingredients taken from free probability theory.

Let us first introduce MOPs.

4.1 Multiple orthogonal polynomials

Let μ be a Borel measure on $\mathbb{R}$ with infinite support and having all its moments. Consider $r \geq 1$ pairwise distinct functions $w_1, \ldots, w_r$ in $L^2(\mu)$.

Definition 4.1. Given a multi-index $\mathbf{n} = (n_1, \ldots, n_r) \in \mathbb{N}^r$, the $\mathbf{n}$-th (type II) MOP associated to the weights $w_1, \ldots, w_r$ and the measure μ is the unique monic polynomial $P_{\mathbf{n}}$ of degree $n_1 + \cdots + n_r$ which satisfies the orthogonality relations

$$\begin{aligned} \int x^k P_{\mathbf{n}}(x) w_1(x) \mu(\mathrm{d}x) &= 0, \qquad 0 \leq k \leq n_1 - 1, \\ &\vdots \qquad\qquad\qquad \vdots \\ \int x^k P_{\mathbf{n}}(x) w_r(x) \mu(\mathrm{d}x) &= 0, \qquad 0 \leq k \leq n_r - 1. \end{aligned} \tag{4.1}$$

Note that the existence/uniqueness of the $\mathbf{n}$-th MOP is not automatic, and depends on whether the system of linear equations (4.1) admits a unique solution. We say that a multi-index $\mathbf{n}$ is normal if it is indeed the case. Since by taking $r = 1$ we clearly recover OPs, we shall assume $r \geq 2$ in what follows.

Let $(\mathbf{n}^{(N)})_{N\in\mathbb{N}} = (n_1^{(N)}, \ldots, n_r^{(N)})_{N\in\mathbb{N}}$ be a sequence of normal multi-indices which satisfies the following path-like structure.

(a) For every $N \in \mathbb{N}$,
$$\sum_{i=1}^{N} n_i^{(N)} = N.$$

(b) For every $N \in \mathbb{N}$ and $1 \leq i \leq r$,
$$n_i^{(N+1)} \geq n_i^{(N)}.$$

(c) There exists $R \in \mathbb{N}$ such that for any $N \in \mathbb{N}$ and $1 \leq i \leq r$,

$$n_i^{(N+R)} \geq n_i^{(N)} + 1. \tag{4.2}$$

(d) For every $1 \leq i \leq r$, there exist $q_1, \ldots, q_r \in (0,1)$ such that

$$\lim_{N \to \infty} \frac{n_i^{(N)}}{N} = q_i. \tag{4.3}$$

We then write for convenience

$$P_N(x) = P_{\mathbf{n}^{(N)}}(x), \qquad N \in \mathbb{N}, \tag{4.4}$$

and observe that P_N has degree N. A question of interest is then to describe the weak convergence of the zero counting probability measure ν_N of P_N as $N \to \infty$, defined as in (1.4) with $z_1, \ldots, z_N$ the zeros of $P_N(x)$, maybe up to a rescaling of the zeros. Before showing how our results answer that question in the case of the multiple Hermite and multiple Laguerre polynomials, we first need to introduce a few ingredients from free probability theory.

4.2 Elements of free probability

Free probability deals with non-commutative random variables which are independent in an algebraic sense. It has been introduced by Voiculescu for the purpose of solving operator algebra problems. We now just provide the few elements of free probability needed for the purpose of this work, and refer to [49, 1] for comprehensive introductions.

For a probability measure λ on $\mathbb{R}$ with compact support, let K_λ be the inverse, for the composition of formal series, of the Cauchy-Stieltjes transform

$$\begin{aligned} G_\lambda(z) &= \int \frac{\lambda(\mathrm{d}x)}{z - x} \\ &= \sum_{k=0}^{\infty} \left(\int x^k \lambda(\mathrm{d}x) \right) z^{-k-1}, \end{aligned} \tag{4.5}$$

and set the R-transform of λ by

$$R_\lambda(z) = K_\lambda(z) - \frac{1}{z}. \tag{4.6}$$

Definition 4.2. Let λ and η be two probability measures on $\mathbb{R}$ with compact support. The free additive convolution of λ and η, denoted by $\lambda \boxplus \eta$, is the unique probability measure (on $\mathbb{R}$ with compact support) which satisfies

$$R_{\lambda \boxplus \eta}(z) = R_\lambda(z) + R_\eta(z). \tag{4.7}$$

Consider a probability measure λ on $[0, +\infty)$ with compact support different from δ_0. If χ_λ is the inverse for the composition of formal series of

$$\frac{1}{z} G_\lambda\left(\frac{1}{z}\right) - 1 = \sum_{k=1}^{\infty} \left(\int x^k \lambda(\mathrm{d}x) \right) z^k, \tag{4.8}$$

we then define the S-transform of λ by

$$S_\lambda(z) = \frac{1+z}{z}\chi_\lambda(z). \tag{4.9}$$

Definition 4.3. Let λ and η be two probability measures on $[0, +\infty)$ with compact support and both different from δ_0. The free multiplicative convolution of λ and η, denoted $\lambda \boxtimes \eta$, is the unique probability measure (on $[0, +\infty)$ with compact support and different from δ_0) which satisfies

$$S_{\lambda\boxtimes\eta}(z) = S_\lambda(z)S_\eta(z). \tag{4.10}$$

For this work, the importance of the free additive and multiplicative convolutions relies on the following results due to Voiculescu, extracted from [1], which describe the limiting eigenvalue distribution of perturbed GUE and Wishart matrices. A random matrix $\mathbf{X}_N$ is distributed according to GUE(N) if it is drawn from the space $\mathcal{H}_N(\mathbb{C})$ of $N \times N$ Hermitian matrices according to the probability distribution

$$\frac{1}{Z_N}\exp\Big\{ - N\mathrm{Tr}(\mathbf{X}_N^2)/2\Big\}\mathrm{d}\mathbf{X}_N, \tag{4.11}$$

where $\mathrm{d}\mathbf{X}_N$ stands for the Lebesgue measure on $\mathcal{H}_N(\mathbb{C}) \simeq \mathbb{R}^{N^2}$ and Z_N is a normalization constant. It is said to be distributed according to Wishart$_\alpha(N)$, where $\alpha \geq 0$ if a real parameter, if the probability distribution reads instead

$$\frac{1}{Z_N}\det(\mathbf{X}_N)^{N\alpha}\exp\Big\{ - N\mathrm{Tr}(\mathbf{X}_N)\Big\}\mathbf{1}_{\big\{\mathbf{X}_N\geq 0\big\}}\mathrm{d}\mathbf{X}_N, \tag{4.12}$$

where $\mathbf{X}_N \geq 0$ means that $\mathbf{X}_N$ is positive semi-definite. The semi-circle distribution is defined by

$$\mu_{\mathrm{SC}}(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbf{1}_{[-2,2]}(x)\mathrm{d}x, \tag{4.13}$$

and the Marchenko-Pastur distribution of parameter $\alpha \geq 0$ by

$$\mu_{\mathrm{MP}(\alpha)}(\mathrm{d}x) = \max(1-\alpha, 0)\delta_0 + \min(\alpha, 1)f_\alpha(x)\mathrm{d}x, \tag{4.14}$$

where, with $\alpha_\pm = (1 \pm \sqrt{\alpha}\,)^2$, we introduced

$$f_\alpha(x) = \frac{1}{2\pi\alpha x}\sqrt{(\alpha_+ - x)(x-\alpha_-)}\,\mathbf{1}_{[\alpha_-,\alpha_+]}(x). \tag{4.15}$$

Then the following holds.

Theorem 4.4. *Consider a sequence of uniformly bounded deterministic matrices $(\mathbf{A}_N)_N$, were $\mathbf{A}_N$ is an $N \times N$ Hermitian matrix, and assume there exists a probability measure λ on $\mathbb{R}$ with compact support such that for all $\ell \in \mathbb{N}$,*

$$\lim_{N\to\infty}\frac{1}{N}\,\mathrm{Tr}\,\big(\mathbf{A}_N\big)^\ell = \int x^\ell\lambda(\mathrm{d}x).$$

(a) *If $(\mathbf{X}_N)_N$ is a sequence of independent random matrices with $\mathbf{X}_N$ distributed according to* GUE(N)*, then for all $\ell \in \mathbb{N}$,*

$$\lim_{N\to\infty}\frac{1}{N}\,\mathbb{E}\Big[\mathrm{Tr}\,\big(\mathbf{X}_N + \mathbf{A}_N\big)^\ell\Big] = \int x^\ell\mu_{\mathrm{SC}}\boxplus\lambda(\mathrm{d}x).$$

(b) *If* $(\mathbf{X}_N)_N$ *is a sequence of independent random matrices with* $\mathbf{X}_N$ *distributed according to* $\mathrm{Wishart}_\alpha(N)$*, and if the* $\boldsymbol{A}_N$*'s are moreover positive semi-definite with* $\lambda \neq \delta_0$*, then for all* $\ell \in \mathbb{N}$*,*

$$\lim_{N\to\infty} \frac{1}{N}\,\mathbb{E}\Big[\,\mathrm{Tr}\left(\mathbf{A}_N^{1/2}\mathbf{X}_N\mathbf{A}_N^{1/2}\right)^{\ell}\Big] = \int x^\ell \mu_{\mathrm{MP}(\alpha)} \boxtimes \lambda(\mathrm{d}x).$$

We are now in position to state the results of this section.

4.3 Multiple Hermite polynomials

Recall that if H_N stands for the N-th Hermite polynomial, that is the OP associated to $\mu(\mathrm{d}x) = e^{-x^2/2}\mathrm{d}x$, then the zero counting probability distribution ν_N of its rescaled version $H_N(\sqrt{N}x)$ is known to converge weakly towards the semi-circle distribution (4.13).

Given $r \geq 2$ pairwise distinct real numbers $a_1, \ldots, a_r$, consider the measure and the weights given by

$$\mu(\mathrm{d}x) = e^{-x^2/2}\mathrm{d}x, \qquad w_j(x) = e^{a_j x}, \qquad 1 \leq j \leq r.$$

The associated MOPs are called multiple Hermite polynomials. For a sequence of multi-indices $(\mathbf{n}^{(N)})_N$ satisfying the path-like structure described in Section 4.1, denote by $H_N^{(a_1,\ldots,a_r)}$ the associated MOP as in (4.4). We shall prove the following.

Theorem 4.5. *Let* ν_N *be the zero probability distribution of the rescaled multiple Hermite polynomial*

$$H_N^{(\sqrt{N}a_1,\ldots,\sqrt{N}a_r)}\big(\sqrt{N}x\big).$$

Then ν_N *converges weakly as* $N \to \infty$ *towards*

$$\mu_{\mathrm{SC}} \boxplus \Big(\sum_{j=1}^r q_j \delta_{a_j}\Big).$$

Although we introduced the R-transform of a probability measure as a formal series, it is actually possible to define it as a proper analytic function, provided one restricts oneself to appropriate subdomains of the complex plane, and equality (4.7) continues to hold, see [6, Section 5]. Then, since $R_{\mu_{\mathrm{SC}}}(z) = z$ and the Cauchy-Stieltjes transform of $\sum_{i=1}^r q_i\delta_{a_i}$ is explicit, one can obtain from (4.7) that the Cauchy-Stiejles transform G of $\mu_{\mathrm{SC}} \boxplus \big(\sum_{j=1}^r q_j\delta_{a_j}\big)$ is an algebraic function, by performing similar manipulations than in the proof of [7, Lemma 1] and concluding by analytic continuation. More precisely, one obtains that G satisfies the algebraic equation

$$P\big(z, G(z)\big) = 0, \qquad z \in \mathbb{C}, \tag{4.16}$$

where the bivariate polynomial $P(z,w)$ is given by

$$P(z,w) = w\prod_{i=1}^r (z - w - a_i) - \sum_{i=1}^r q_i \prod_{j=1,\, j\neq i}^r (z - w - a_i). \tag{4.17}$$

Probability measures for which the Cauchy-Stieltjes transform is algebraic have interesting regularity properties, see [2, Section 2.8], and are moreover suitable for numerical evaluation, see e.g. [25].

We now turn to multiple Laguerre polynomials, for which we provide a similar analysis.

4.4 Multiple Laguerre polynomials

If $L_N^{(\alpha)}$ stands for the N-th Laguerre polynomial of parameter $\alpha \geq 0$, that is the OP associated to $\mu(\mathrm{d}x) = x^\alpha e^{-x}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x$, then it is known that the zero probability distribution ν_N of $L_N^{(N\alpha)}(Nx)$ converges weakly as $N \to \infty$ towards the Marchenko-Pastur distribution of parameter α (4.14)–(4.15).

There exist two different definitions for the multiple Laguerre polynomials in the literature, see [30, Section 23.4]. We consider here the so-called multiple Laguerre polynomials of the second kind, which are defined as follows. Given $r \geq 2$ pairwise distinct positive numbers $a_1, \ldots, a_r$ and $\alpha \geq 0$, consider

$$\mu(\mathrm{d}x) = x^\alpha e^{-x}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x, \qquad w_j(x) = e^{(1+a_j)x}, \qquad 1 \leq j \leq r,$$

and, given a sequence of multi-indices $(\mathbf{n}^{(N)})_N$ satisfying the path-like structure described previously, let $L_N^{(\alpha\,;\,a_1\,,\,\ldots\,,\,a_r)}$ be the associated MOP as in (4.4).

Theorem 4.6. *Let ν_N be the zero probability distribution of the rescaled multiple Laguerre polynomial*

$$L_N^{(N\alpha\,;\,Na_1\,,\,\ldots\,,\,Na_r)}(x).$$

Then ν_N converges weakly as $N \to \infty$ towards

$$\mu_{\mathrm{MP}(\alpha)} \boxtimes \Big(\sum_{j=1}^r q_j \delta_{1/a_j}\Big).$$

As it was the case for the R-transform, the S-transform can be defined as an analytic function, and (4.10) also holds on subdomains of the complex plane, see [6, Section 6]. Then, because $S_{\mu_{\mathrm{MP}(\alpha)}}(z) = (1+\alpha z)^{-1}$, one can also obtain from (4.10), taking care of the definition domains, that the Cauchy-Stieltjes transform G of $\mu_{\mathrm{MP}(\alpha)} \boxtimes \big(\sum_{j=1}^r q_j\delta_{1/a_j}\big)$ satisfies the algebraic equation

$$P\big(z, G(z)\big) = 0, \qquad z \in \mathbb{C}, \tag{4.18}$$

where $P(z,w)$ is given by

$$P(z,w) = w\prod_{i=1}^r \big(z - \frac{1}{a_i}(1-\alpha+\alpha z w)\big) - \sum_{i=1}^r q_i \prod_{j=1,\, j\neq i}^r \big(z - \frac{1}{a_i}(1-\alpha+\alpha z w)\big). \tag{4.19}$$

4.5 Proofs

Before providing proofs for Theorems 4.5 and 4.6, we first precise a few points concerning MOP Ensembles, that we introduced in Section 1.3.

A sequence of measures $(\mu_N)_N$, weights $w_{j,N} \in L^2(\mu_N)$, $1 \le j \le r$, and a path-like sequence of multi-indices $(\mathbf{n}^{(N)})_N$ induce a sequence of MOP Ensembles. Namely, for each N one can associate random variables $x_1, \dots, x_N$ distributed according to (1.7) where we chose for the multi-index $\mathbf{n} = \mathbf{n}^{(N)}$. For the monic polynomial $P_{k,N}$ of degree k appearing in the kernel (1.9), one can choose the $\mathbf{n}^{(k)}$-th (type II) MOP associated with μ_N and the $w_{j,N}$'s. The associated biorthogonal functions $Q_{k,N}$'s can then be constructed from the type I MOPs, see [30, Theorem 23.1.6], and Assumption 1.1 is satisfied with $\mathfrak{q}_N = 1$. We moreover recall that the average characteristic polynomial χ_N equals $P_{N,N}$.

In order to obtain growth estimates for the $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s, we now describe a connection with the so-called nearest neighbors recurrence coefficients, which are in practice easier to compute.

4.5.1 Nearest neighbors recurrence coefficients

Van Assche [48] established for general MOPs, says associated to a measure μ and weights w_i's, that for every normal multi-index $\mathbf{n}$ there exist real numbers $(a_{\mathbf{n}}^{(d)})_{1\le d\le r}$ and $(b_{\mathbf{n}}^{(d)})_{1\le d\le r}$ satisfying

$$\begin{aligned} xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_1} + a_{\mathbf{n}}^{(1)} P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)} P_{\mathbf{n}-\mathbf{e}_d}(x), \\ &\vdots \\ xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_r} + a_{\mathbf{n}}^{(r)} P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)} P_{\mathbf{n}-\mathbf{e}_d}(x), \end{aligned} \tag{4.20}$$

where

$$\mathbf{e}_d = (\underbrace{0,\dots,0}_{d-1},1,0,\dots,0) \in \mathbb{N}^r, \qquad 1 \le d \le r.$$

Note that this provides

$$P_{\mathbf{n}+\mathbf{e}_i}(x) - P_{\mathbf{n}+\mathbf{e}_j}(x) = (a_{\mathbf{n}}^{(j)} - a_{\mathbf{n}}^{(i)}) P_{\mathbf{n}}(x), \qquad 1 \le i,j \le r. \tag{4.21}$$

With the path-like sequence of multi-indices $(\mathbf{n}^{(k)})_{k\in\mathbb{N}}$ and allowing the w_i's and μ to depend on a parameter N, we write for convenience

$$a_{k,N}^{(d)} = a_{\mathbf{n}^{(k)},N}^{(d)}, \qquad b_{k,N}^{(d)} = b_{\mathbf{n}^{(k)},N}^{(d)}, \qquad 1 \le d \le r.$$

Then the following holds.

Lemma 4.7. *If there exists $\varepsilon > 0$ such that for every $1 \le d \le r$ the sequences*

$$\left\{ \max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\le\varepsilon} \max_{j=1}^{r} \left| a_{\mathbf{n}^{(k)}-\mathbf{e}_j,N}^{(d)} \right| \right\}_{N\ge1}, \qquad \left\{ \max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\le\varepsilon} \left| b_{k,N}^{(d)} \right| \right\}_{N\ge1}, \tag{4.22}$$

are bounded, then so is the sequence

$$\left\{ \max_{k,m\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\le\varepsilon,\,\left|\frac{m}{N}-1\right|\le\varepsilon} \left| \langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} \right| \right\}_{N\ge1}.$$

Proof. First, as a consequence of (4.2) and [30, (23.1.7)], we have

$$\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} = 0, \qquad m < k - R. \tag{4.23}$$

Define the sequence $(i_k)_{k\in\mathbb{N}}$ taking its values in $\{1,\ldots,r\}$ by

$$\mathbf{n}^{(k+1)} = \mathbf{n}^{(k)} + \mathbf{e}_{i_k}, \qquad m \in \mathbb{N}.$$

For a fixed k, which may be chosen as large as we want, (4.20) yields

$$xP_{k,N}(x) = P_{k+1,N}(x) + a_{k,N}^{(i_k)} P_{k,N}(x) + \sum_{d=1}^{r} b_{k,N}^{(d)} P_{\mathbf{n}^{(k)}-\mathbf{e}_d,N}(x). \tag{4.24}$$

Then, since (4.21) provides for any $1 \le d \le r$ and m large enough

$$P_{\mathbf{n}^{(m)}-\mathbf{e}_d,N}(x) = P_{m-1,N}(x) + \big(a^{(d)}_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N} - a^{(i_{m-1})}_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N}\big) P_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N}(x),$$

we obtain inductively with (4.24) that

$$\begin{aligned} xP_{k,N}(x) &= P_{k+1,N}(x) + a_{k,N}^{(i_k)} P_{k,N}(x) + \Big(\sum_{d=1}^{r} b_{k,N}^{(d)}\Big) P_{k-1,N}(x) \\ &+ \sum_{m=k-R}^{k-2} \left(\sum_{d=1}^{r} b_{k,N}^{(d)} \prod_{l=m+1}^{k-1} \big(a^{(d)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N} - a^{(i_l)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N}\big) \right) P_{m,N}(x) + R_{k,N}(x), \end{aligned} \tag{4.25}$$

where $R_{k,N}$ is a polynomial of degree at most $k-R-1$. By comparing (4.25) with the (unique) decomposition (1.18) and (4.23), we obtain explicit formulas for the $\langle xP_{k,N}, Q_{m,N}\rangle$'s in terms of the nearest neighbor recurrence coefficients, from which Lemma 4.7 easily follows. $\square$

4.5.2 Proof of Theorem 4.5

Proof. Associate to the multi-indices $(\mathbf{n}^{(N)})_{N\in\mathbb{N}}$ the (uniformly bounded) sequence $(\mathbf{A}_N)_{N\in\mathbb{N}}$ of diagonal matrices

$$\mathbf{A}_N = \mathrm{diag}\big(\underbrace{a_1,\ldots,a_1}_{n_1^{(N)}}, \ldots, \underbrace{a_r,\ldots,a_r}_{n_r^{(N)}}\big) \in \mathcal{H}_N(\mathbb{C}).$$

On the one hand, let $(\mathbf{X}_N)_N$ be a sequence of independent random matrices, with $\mathbf{X}_N$ distributed according to GUE(N). If $\hat\mu^N$ stands for the empirical measure associated to the eigenvalues of $\mathbf{Y}_N = \mathbf{X}_N + \mathbf{A}_N$, then Theorem 4.4 (a) and (4.3) provide for any $\ell \in \mathbb{N}$

$$\begin{aligned} \lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat\mu^N(\mathrm{d}x)\right] &= \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\Big[\mathrm{Tr}\left(\mathbf{X}_N + \mathbf{A}_N\right)^\ell\Big] \\ &= \int x^\ell \mu_{\mathrm{SC}} \boxplus \Big(\sum_{j=1}^r q_j \delta_{a_j}\Big)(\mathrm{d}x). \end{aligned} \tag{4.26}$$

On the other hand, observe from (4.11) that the random matrix $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z'_N}\exp\Big\{-N\mathrm{Tr}\big(\mathbf{Y}_N^2-2\mathbf{A}_N\mathbf{Y}_N\big)/2\Big\}\mathrm{d}\mathbf{Y}_N, \tag{4.27}$$

where Z'_N is a new normalization constant. By performing a spectral decomposition in (4.27), integrating out the eigenvectors and using a confluent version of the Harish-Chandra-Itzykson-Zuber formula, Bleher and Kuijlaars [9] obtained that the random eigenvalues of $\mathbf{Y}_N$ form a MOP Ensemble, see (1.7), associated to the N-dependent weights and measure

$$\mu_N(\mathrm{d}x)=e^{-Nx^2/2}\mathrm{d}x, \qquad w_{j,N}(x)=e^{Na_jx}, \qquad 1\le j\le r, \tag{4.28}$$

and the multi-index $\mathbf{n}^{(N)}$. The average characteristic polynomial χ_N for that MOP Ensemble then equals the associated $\mathbf{n}^{(N)}$-th MOP, which is seen from a change of variable to be $H_N^{(\sqrt{N}a_1,\dots,\sqrt{N}a_r)}(\sqrt{N}x)$, up to a multiplicative constant. The weights in (4.28) form an AT system, from which it follows that any multi-index is normal, and that χ_N has real zeros, cf. [30, Chapter 23]. One moreover obtains from [48, Section 5.2] and a change of variables explicit formulas for the nearest neighbors recurrence coefficients associated to (4.28),

$$a^{(d)}_{\mathbf{n},N}=a_d, \qquad b^{(d)}_{\mathbf{n},N}=\frac{n_d}{N}, \qquad \mathbf{n}=(n_1,\dots,n_r).$$

Thus, Theorem 4.5 follows from (4.26), Lemma 4.7 and Corollary 1.6. □

4.5.3 Proof of Theorem 4.6

Proof. The proof follows the same spirit as the proof of Theorem 4.5. Introduce the sequence of (uniformly bounded) diagonal matrices

$$\mathbf{A}_N=\mathrm{diag}\big(\underbrace{1/a_1,\dots,1/a_1}_{n_1^{(N)}},\dots,\underbrace{1/a_r,\dots,1/a_r}_{n_r^{(N)}}\big),$$

and let $(\mathbf{X}_N)_N$ be a sequence of independent random matrices, where $\mathbf{X}_N$ is distributed according to $\mathrm{Wishart}_\alpha(N)$. With $\hat{\mu}^N$ the empirical measure of the eigenvalues of $\mathbf{Y}_N=\mathbf{A}_N^{1/2}\mathbf{X}_N\mathbf{A}_N^{1/2}$, Theorem 4.4 (b) and (4.3) then provide for all $\ell\in\mathbb{N}$

$$\lim_{N\to\infty}\mathbb{E}\left[\int x^\ell\hat{\mu}^N(\mathrm{d}x)\right]=\int x^\ell\mu_{\mathrm{MP}(\alpha)}\boxtimes\Big(\sum_{j=1}^r q_j\delta_{1/a_j}\Big)(\mathrm{d}x). \tag{4.29}$$

Now, observe from (4.12) that $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z'_N}\det(\mathbf{Y}_N)^{N\alpha}\exp\Big\{-N\mathrm{Tr}\big(\mathbf{A}_N^{-1}\mathbf{Y}_N\big)\Big\}\mathbf{1}_{\{\mathbf{Y}_N\ge 0\}}\mathrm{d}\mathbf{Y}_N, \tag{4.30}$$

where Z'_N is a new normalization constant. Similarly than for the Hermite case, the eigenvalues of $\mathbf{Y}_N$ form a MOP Ensemble associated to

$$\mu_N(\mathrm{d}x)=x^{N\alpha}e^{-Nx}\mathrm{d}x, \qquad w_{j,N}(x)=e^{N(1+a_j)x}, \qquad 1\le j\le r, \tag{4.31}$$

and the multi-index $\mathbf{n}^{(N)}$, see [10]. The average characteristic polynomial χ_N is then the $\mathbf{n}^{(N)}$-th MOP associated to (4.31), which is $L_N^{(N\alpha\,;\,Na_1\,,\,\ldots\,,\,Na_r)}(x)$ up to a multiplicative constant. The weights in (4.31) form an AT system so that any multi-index is normal and χ_N has real zeros. If we denotes $|\mathbf{n}| = n_1 + \cdots + n_r$ for $\mathbf{n} \in \mathbb{N}^r$, then one obtains from [48, Section 5.4] that the nearest neighbors recurrence coefficients for (4.28) read

$$a^{(d)}_{\mathbf{n},N} = \frac{n_d(|\mathbf{n}| + N\alpha)}{N^2 a_d}, \qquad b^{(d)}_{\mathbf{n},N} = \frac{|\mathbf{n}| + N\alpha + 1}{N a_d} + \sum_{j=1}^{r} \frac{n_j}{N a_j}, \qquad \mathbf{n} = (n_1, \ldots, n_r).$$

Theorem 4.6 finally follows from (4.29), Lemma 4.7 and Corollary 1.6. □

Remark 4.8. Having in mind the proofs of Theorems 4.5 and 4.6, it would be of interest to find out if there exists a matrix model for the multiple version of the Jacobi polynomials, the Jacobi-Piñeiro polynomials, which are related in a limiting case to the rational approximations of $\zeta(k)$ and polylogarithms [30, Section 23.3.2], and then if it would be possible to describe its limiting zero distribution thanks to free convolutions.

Acknowledgements The author would like to thank Walter Van Assche for interesting discussions and the references [42, 46]. He is also grateful to Steven Delvaux for pointing out the relation (4.21) and for the interesting discussion which followed.

The author is supported by FWO-Flanders projects G.0427.09 and by the Belgian Interuniversity Attraction Pole P07/10.

References

[1] G. W. Anderson, A. Guionnet and O. Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics Vol. 118, Cambridge University Press, Cambridge, U.K. (2010).

[2] G. W. Anderson and O. Zeitouni, *A law of large numbers for finite-range dependent random matrices*, Comm. Pure Appl. Math. 61 (2008), 1118–1154.

[3] J. Arvesú, J. Coussement and W. Van Assche, *Some discrete multiple orthogonal polynomials*, J. Comput. Appl. Math 153 (2003), 19–45.

[4] A. I. Aptekarev, A. Branquinho and W. Van Assche, *Multiple orthogonal polynomials for classical weights*, Trans. Amer. Math. Soc. 355 (2003), 3887–3914.

[5] M. Bender, S. Delvaux and A. B. J. Kuijlaars, *Multiple Meixner-Pollaczek polynomials and the six-vertex model*, J. Approx. Theory 163 (2011), 1606–1637.

[6] H. Bercovici and D. Voiculescu, *Free convolution of measures with unbounded support*, Indiana Univ. Math. J. 42 (1993), 733–773.

[7] P. Biane, *On the free convolution with a semi-circular distribution*, Indiana Univ. Math. J. 46 (1997), 705–718.

[8] P. M. Bleher, S. Delvaux and A. B. J. Kuijlaars, *Random matrix model with external source and a constrained vector equilibrium problem*, Comm. Pure Appl. Math. 64 (2011), no. 1, 116–160.

[9] P. M. Bleher and A. B. J. Kuijlaars, *Random matrices with external source and multiple orthogonal polynomials*, Int. Math. Res. Not. (2004), 109–129.

[10] P. M. Bleher and A. B. J. Kuijlaars, *Integral representations for multiple Hermite and multiple Laguerre polynomials*, Ann. Inst. Fourier 55 (2005), 2001–2014.

[11] A. Borodin, *Biorthogonal ensembles*, Nuclear Phys. B 536 (1999), no. 3, 704–732.

[12] A. Borodin and G. Olshanski, *Meixner polynomials and random partitions*, Moscow Math. J. 6 (2006), 629–655.

[13] B. Collins, *Product of random projections, Jacobi ensembles and universality problems arising from free probability*, Probab. Theory Related Fields 133 (2005), 315–344.

[14] E. Coussement and W. Van Assche, *Some classical multiple orthogonal polynomials*, J. Comput. Appl. Math. 127 (2001), 317–347.

[15] E. Daems and A. B. J. Kuijlaars, *Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions*, J. Approx. Theory 146 (2007), 91–114.

[16] P. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics Vol. 3, AMS Providence R.I. (1999).

[17] S. Delvaux, *Average characteristic polynomials for multiple orthogonal polynomial ensembles*, J. Approx. Theory 162 (2010), 1033–1067.

[18] P. Desrosiers and P. J. Forrester, *A note on biorthogonal ensembles*, J. Approx. Theory 152(2) (2008), 167–187.

[19] H. Dette and L. A. Imhof, *Uniform approximation of eigenvalues in Laguerre and Hermite β-ensembles by roots of orthogonal polynomials*, Trans. Amer. Math. Soc. 359 (2007), 4999–5018.

[20] H. Dette and J. Nagel, *Some asymptotic properties of the spectrum of the Jacobi ensemble*, SIAM J. Math. Anal. 41 (2009), 1491–1507.

[21] M. Duits, D. Geudens and A. B. J. Kuijlaars, *A vector equilibrium problem for the two-matrix model in the quartic/quadratic case*, Nonlinearity 24 (2011), 951–993.

[22] M. Duits and A. B. J. Kuijlaars, *Universality in the two matrix model: a Riemann-Hilbert steepest descent analysis*, Comm. Pure Appl. Math. 62 (2009), 1076–1153.

[23] M. Duits, A. B. J. Kuijlaars and M. Y. Mo, *The Hermitian two-matrix model with an even quartic potential*, Memoirs Amer. Math. Soc. 217 No. 1022 (2012), 1–105.

[24] M. Duits, A. B. J. Kuijlaars and M.Y. Mo, *Asymptotic analysis of the two matrix model with a quartic potential*, Preprint arXiv:1210.0097.

[25] A. Edelman and N. R. Rao, *The polynomial method for random matrices*, Found. Comput. Math. 8 (2008), 649–702.

[26] W. Van Assche, J. S. Geronimo and A. B. J. Kuijlaars, *Riemann–Hilbert problems for multiple orthogonal polynomials*, pp. 23–59 in: "NATO ASI Special Functions 2000. Current Perspective and Future Directions", (J. Bustoz, M.E.H. Ismail, S.K. Suslov eds.), Nato Science Series II Vol.30, Kluwer Academic Publishers, Dordrecht, 2001.

[27] I. Gohberg, S. Goldberg, and N. Krupnik, *Traces and Determinants of linear Operators*, Birkhäuser, Basel (2000).

[28] A. Hardy and A. B. J. Kuijlaars, *Large deviations for a non-centered Wishart matrix*, preprint arXiv:1204.6261 (2012).

[29] J. B. Hough, M. Krishnapur, Y. Peres and B. Virág, *Determinantal processes and independence*, Probab. Surveys Volume 3 (2006), 206–229.

[30] M. E. H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, Encyclopedia of Mathematics and its Applications 98, Cambridge University Press, 2005.

[31] K. Johansson, *Random matrices and determinantal processes*, Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005 (Bovier et al., eds.), Elsevier, 2006, 1–55.

[32] K. Johansson, *Shape fluctuations and random matrices*, Commun. Math. Phys. 209 (2000), 437–476.

[33] K. Johansson, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, Annals of Mathematics 153 (2001), 259–296.

[34] R. Koekoek, P. A. Lesky and R. F. Swarttouw, *Hypergeometric orthogonal polynomials and their q-analogues*. Springer Monographs in Mathematics. Springer-Verlag, Berlin (2010).

[35] W. König, *Orthogonal polynomial ensembles in probability theory*, Probab. Surveys Volume 2 (2005), 385–447.

[36] A. B. J. Kuijlaars, *Multiple orthogonal polynomial ensembles*, Recent trends in orthogonal polynomials and approximation theory, 155–176, Contemp. Math., 507, Amer. Math. Soc., Providence, RI (2010).

[37] A. B. J. Kuijlaars, *Multiple orthogonal polynomials in random matrix theory*, Proceedings of the International Congress of Mathematicians. Volume III, 1417–1432, Hindustan Book Agency, New Delhi (2010).

[38] A. B. J. Kuijlaars and K. T.-R. McLaughlin, *A Riemann-Hilbert problem for biorthogonal polynomials*, J. Comput. Appl. Math. 178 (2005), 313–320.

[39] A. B. J. Kuijlaars, A. Martínez-Finkelshtein and F. Wielonsky, *Non-intersecting squared Bessel paths and multiple orthogonal polynomials for modified Bessel weights*, Comm. Math. Phys. 286 (2009), 217–275.

[40] A. B. J. Kuijlaars and P. Román, *Recurrence relations and vector equilibrium problems arising from a model of non-intersecting squared Bessel paths*, J. Approx. Theory 162 (2010), 2048–2077.

[41] A. B. J. Kuijlaars and W. Van Assche, *The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients*, J. Approx. Theory 99 (1999), 167–97.

[42] P. Nevai, *Orthogonal polynomials*, Mem. Amer. Math. Soc. 213 (1979).

[43] B. Simon, *Weak convergence of CD kernels and applications*, Duke Math. J. Volume 146 (2009), 305–330.

[44] B. Simon, *Trace ideals and their applications*, volume 120 of *Mathematical Surveys and Monographs.* Providence, RI, American Mathematical Society, second edition, 2005.

[45] A. Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys 55 (2000), 923–975.

[46] W. Van Assche, *Eigenvalues of Toeplitz matrices associated with orthogonal polynomials*, J. Approx. Theory 51 (1987) 360–371.

[47] W. Van Assche, *Multiple orthogonal polynomials, irrationality and transcendence*, Continued fractions: from analytic number theory to constructive approximation (University of Missouri–Columbia, Columbia, MO, USA 1998), Contemp. Math., vol. 236, Amer. Math. Soc., Providence, RI 1999, pp. 325–342.

[48] W. Van Assche, *Nearest neighbor recurrence relations for multiple orthogonal polynomials*, J. Approx. Theory 163 (2011), 1427–1448.

[49] D. V. Voiculescu, K. J. Dykema and A. Nica, *Free random variables*, CRM Monographs, vol. 1, American Math. Soc. (1992).